

# Logical Device Cache Process [E]

---

Logical device cache is an optional feature that you may enable to reduce disk I/O wait time from a user's perspective. On CRAY J90 systems, logical cache is defined in central memory (DDRAM).

When a process issues a read request of data on a file system, the action taken to access the data depends on whether the data is currently in the UNICOS system buffer cache or logical device cache. The process is described as follows:

1. If the data is found in the system buffer cache (central memory), it is copied to the user area. If the requested data is not found, step 2 is taken.
2. If ldcache (logical device cache) has been allocated for the file system, the ldcache area is searched for the sector of data. If found, it is read into the system buffer cache and then copied to the user's process space. If the desired data is not found, step 3 is taken.
3. If ldcache is allocated for the file system, the sector is read from disk and cached into the ldcache area. The sector is then read from the ldcache device into the system buffer (central memory) cache and then copied to the user area.

**Note:** The system buffer cache may be bypassed if the data is a multiple of 512 words, begins on a word boundary, and the file system address of the data is on a block boundary.

The system buffer cache writes only to the ldcache area. When system buffers age and require reassignment, the system buffers are written to ldcache and the ldcache segment is marked as dirty. Dirty segments in ldcache are then written to disk when the segment is needed for a different part of the file system, when the ldcache area is flushed to disk by `ldsync(8)`, or when the system periodically flushes the ldcache area to disk.

## E.1 Setting up ldcache by using `/etc/ldcache`

The cache for a logical device is specified as a number of units and a count of 4096-byte blocks per unit. The system administrator easily configures the relationship between the number of cache units and the cache unit size. The `/tmp` and `root (/)` file systems are excellent candidates for logical device cache. If you have more ldcache area available, distribute the remaining area to other heavily used file systems. To be effective, the ldcache hit rate should be

above 97% for ld caching; however, the main concern is the ratio of logical reads to physical reads.

The `/etc/ldcache` command assigns groups of blocks, called *units*, of an ld cache device (central memory) to a specific file system. To set the number of blocks in an ld cache unit, use the `ldcache -s` command. Choose the size that is used in the `mkfs` command to build that file system. This makes reads and writes to that physical device much faster.

If a striped file system is cached, multiply the number of blocks per cylinder for the physical device type by the number of devices in the stripe group. Larger unit sizes are good for sequential I/O, but they may cause excessive I/O when the I/O is random.

Ensure that the number of blocks assigned for ld cache for all file systems added together does not exceed the total number of blocks available on your logical cache device. To calculate this figure, use the following steps:

1. For each file system being ld cached, multiply the number of blocks in an ld cache unit by the total number of ld cache units allocated for that file system.
2. Add all such totals together.
3. Subtract that sum from the total number of blocks available on the ld cache device for ld caching.

## E.2 Assigning ld cache

When assigning logical device cache, be sure to include the type. The `MEM` type is used when assigning central memory-based logical device cache. The `LDCHCORE` value defines the number of blocks of core memory to be used for logical device cache. The configuration specification language (CSL) `NLDCH` value defines the number of cache headers that will be configured. This sets the total number of logical device cache units that can be active at one time. You must use both the `CSL LDCHCORE` and `NLDCH` statements in conjunction to define central, memory-based logical device cache.

```
/etc/ldcache -l dev -n units [-s size] [-t type]
```

`-l dev`

Full path name or minor device number of logical device.

`-n units`

Number of cache units to assign. If 0, the logical device caching is released.

**-s *size***                      Size (in 4-Kbyte blocks) of each cache unit. For best performance, set *size* as a multiple of tracks per cylinder related to the logical device and the file system used.

**-t *type***                      Type of memory for cache (MEM).

An example of releasing a logical device cache follows:

```
# /etc/ldcache -l /dev/dsk/user_a -n 0
```

An example of assigning a logical device cache follows:

```
# /etc/ldcache -l /dev/dsk/source-tree -s 27 -t MEM -n 500
```

You also can assign a logical device cache by creating an `/etc/config/ldchlist` file, which contains logical device cache configuration information used by `/etc/rc`. During multiuser startup, the `/etc/rc` script checks for the existence of an `/etc/config/ldchlist` file. If the file exists, `/etc/rc` will configure ld caching according to the entries and values in the `/etc/config/ldchlist` file. There are four fields per line, separated by space; the first field is the logical device, the second field is the cache type (MEM), the third field is the number of cache units, and the fourth field is the size in 4-Kbyte blocks of each unit (usually a track size). The following is an example:

```
/dev/dsk/root MEM 300 27
/dev/dsk/usr MEM 300 27
/dev/dsk/tmp MEM 300 27
/dev/dsk/home MEM 300 27
```

The third field multiplied by the fourth field is the total cache area (in blocks) allocated for that file system. The total of the third column is the number of NLDCH that you must define in the UNICOS `config` file.

An example of displaying the ldcache hit rate follows:

```
# /etc/ldcache
T unit size      reads      writes      hits      misses      rate      name
- - - - -
B 300 27      16727      30354      34865      1799      95.09      /dev/dsk/root
B 300 27      1729       4703       1399       254       84.63      /dev/dsk/home
B 250 27      6702       20794      6191       263       95.93      /dev/dsk/tmp
#
M 200 10       47         11         27         4         87.10      /dev/dsk/src

# ldcache -bCache to user      Cache to disk      Cache/disk ratio
Reads  Writes      Reads  Writes      Read  Write  Total  Name
-----
839155 334505      65016 28772      12.9 11.6  12.5   /dev/dsk/root_b
301039 26871       25616 1628       11.8 16.5  12.0   /dev/dsk/usr_b
 68947 74725       13424 13824       5.1  5.4   5.3   /dev/dsk/spool
   183 1678       18416 1743       0.0  1.0   0.1   /dev/dsk/usr_tmp

# ldcache -l /dev/dsk/tmp /dev/dsk/tmp  Fri Sep 24 14:52:12 1993
```

A hit rate of under 97% probably indicates that the file system is not a good candidate for ld caching or that you should enlarge the size of that file system's ld cache area if possible. In the preceding display, you should examine the file system usage and ld cache configuration aspects of the /dev/dsk/home and /dev/dsk/src file systems. You also should examine the ratio of logical reads to physical reads, as shown in the preceding display of the ld cache -b example.

An example of displaying ld cache statistics for an individual file system follows:

```

Read data      Write data
-----
Blocks transferred:      689      1296
Avg request length:      1 blks      1 blks
Lst transfer rate:      0.008192 Mbs      0.061236 Mbs
Max transfer rate:      0.135680 Mbs      0.208438 Mbs
Cache hits:      597      677
Cache misses:      0      73
Cache hit rate:      1000.000000      90.266667
```

### E.3 Flushing data by using `/etc/ldsync`

You can use the `/etc/ldsync` command to flush data from all logical device caches to disk. Only data that has been written to a logical device cache, but not to disk, is affected. The `/etc/ldsync` command does not flush data in the system buffers to disk. During normal operation, the UNICOS system periodically flushes data from the `ldcache` area to disk; the `/bin/sync` command does this action.

During a normal UNICOS system shutdown, all logical device cache data is flushed to disk. At shutdown time it is important that all `ldcache` is removed from all file systems. To check that all `ldcache` is removed, use `/etc/ldcache`. The command should print just a header, as in the following example:

```
# /etc/ldcache

T  Unit  Size  Reads  Writes  Hits  Misses  Rate  Name
-----
#
```

For additional information about when to execute the `/etc/ldsync` command when shutting down the UNICOS system, see the procedure in Chapter 3, page 19.

